MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LINCOLN LABORATORY

PACKET SPEECH SYSTEMS TECHNOLOGY

SEMIANNUAL TECHNICAL SUMMARY REPORT TO THE DEFENSE ADVANCED RESEARCH PROJECTS AGENCY

1 APRIL — 30 SEPTEMBER 1982

ISSUED 3 FEBRUARY 1983

Approved for public release; distribution unlimited.

LEXINGTON MASSACHUSETTS

ABSTRACT

This report describes work performed on the Packet Speech Systems Technology Program sponsored by the Information Processing Techniques Office of the Defense Advanced Research Projects Agency during the period 1 April through 30 September 1982.

TABLE OF CONTENTS

	Abstract	iii
	Introduction and Summary	vii
I.	COMPACT LPC VOCODER DEVELOPMENT	1
II.	COMPACT EMBEDDED VOCODER	3
III.	PACKET VOICE TERMINAL	15
	A. PVT and LEXNET Hardware Status	15
	B. PVT Technology Transfer	15
	C. PVT Software Status	16
	D. Measurement Host Development	18
IV.	MINICONCENTRATOR	21
	A. Hardware Status	21
	B. Gateway Software	21
	1. Gateway Program	21
	2. Support Software	23
V.	WIDEBAND NETWORK EXPERIMENTS AND EXPERIMENT COORDINATION	25
VI.	VOICE CONTROL OF NETWORK VOICE CONFERENCING	35
	A. Overview	35
	B. Recognition Accuracy over a LEXNET	35
	C. Performance of Human/VCOP Interaction During Conference Setup	37
	D. VCOP/NVP Interaction	38
	E. Network Tests	38
References		40

INTRODUCTION AND SUMMARY

The long-range objectives of the Packet Speech Systems Technology Program are to develop and demonstrate techniques for efficient digital speech communications on networks suitable for both voice and data, and to investigate and develop techniques for integrated voice and data communication in packetized networks, including wideband common-user satellite links. Specific areas of concern are: the concentration of statistically fluctuating volumes of voice traffic, the adaptation of communication strategies to varying conditions of network links and traffic volume, and the interconnection of wideband satellite networks to terrestrial systems.

Previous efforts in this area have led to new vocoder structures for improved narrowband voice performance and multiple-rate transmission, and to demonstrations of conversational speech and conferencing on the ARPANET and the Atlantic Packet Satellite Network.

The current program has two major thrusts: i.e., the development and refinement of practical low-cost, robust, narrowband, and variable-rate speech algorithms and voice terminal structures; and the establishment of an experimental wideband satellite network to serve as a unique facility for the realistic investigation of voice/data networking strategies.

This report covers work in the following areas: compact vocoder development based on digital LSI technology; development of Packet Voice Terminal (PVT) and local access network (LEXNET) facilities for experiments in packet voice; development and experimental application of a flexible internet stream gateway (the miniconcentrator) for packet voice; planning, coordination, and execution of multi-user packet speech experiments on the experimental wideband satellite network (WB SATNET); and design and development of a system for voice control of network voice conferencing in the wideband system.

Ten single-card 2400-bps LPC vocoders, using factory-programmed versions of the NEC Signal-Processing Interface (SPI) chips, have been constructed, debugged, and are operational in PVTs in the experimental wideband network.

A compact embedded LPC vocoder system which operates in the 800- to 4800-bps range, and can be realized in hardware using the NEC SPI chips as programmed for the 2400-bps compact LPC, has been developed. Both five-rate and three-rate systems have been implemented in real time on the Lincoln Digital Signal Processor (LDSP). Architectures and microcode have been developed for hardware realizations using the NEC SPIs and an INTEL 8085 control and interface microcomputer.

LEXNETs and PVTs are operating at four wideband network sites and have been used in a variety of experiments and demonstrations. A draft specification has been written for production of PVTs by an industrial contractor. PVT conferencing software, including robustness features and three methods of conference setup, is operational. A measurement host capability, including a timer card to allow precise cross-net packet delay measurements based on a global time source, has been added to the PVT.

Miniconcentrator gateway hardware has been running reliably at four sites throughout this reporting period. Several major extensions to the PDP-11 gateway program have been added, in the areas of speech conferencing, robustness, measurement and instrumentation facilities, and external control of the gateway. The gateways were a major element in demonstrations of internet packet speech and conferencing, as well as in a large set of experiments and performance validations on the wideband system.

Lincoln carried out experiment planning and coordination, as well as development of PVTs, LEXNETs, and gateways, for a 3 June demonstration at the final DARPA narrowband packet speech program meeting. The demonstration, which in itself was a major experimental milestone for the wideband system, involved the Lincoln, ISI, and SRI sites. Point-to-point and conference calls using 2400-bps LPC, 16- to 64-kbps embedded CVSD, and 64-kbps PCM were demonstrated. Calls of particular interest included: (1) a PCM call from a LEXNET PVT at Lincoln and to the local Los Angeles weather number through a switched telephone network interface (STNI) at ISI; (2) a 3-site, 4-party LPC conference using single-card compact LPCs in PVTs on LEXNETs at Lincoln and ISI, and a CHI-V LPC on a packet radio net (PRNET) unit at ISI; and (3) a call between an LPC on a Lincoln LEXNET, and an SRI LPC in a mobile van

PRNET unit, demonstrating mobile PRNET speech and alternate packet routing as the van traveled in the Palo Alto area.

The voice-controlled operator (VCOP) has been implemented and tested. LPC and PCM conferences involving two to four participants have been successfully set up by voice between two LEXNETs connected through a gateway. The conference initiator sets up the conference by calling VCOP from any PVT and carrying on a voice dialogue structured by means of voice prompts from VCOP.

PACKET SPEECH SYSTEMS TECHNOLOGY

I. COMPACT LPC VOCODER DEVELOPMENT

In the previous semiannual, it was reported that a pair of single-card 2.4-kbps PVT-compatible LPC vocoders had been demonstrated using EPROM versions of the Signal-Processing Interface (SPI) chips (μPD7720). During the period, an order was placed with NEC Electronics, U.S.A., for 150 factory-programmed ROM versions of each of the three SPI single-chip microcomputers used in the single-card LPC vocoder. The order has been received in full and checks on approximately one-third of the 450 pieces have found them to be functional. A total of 14 PVT-compatible LPC boards have been fabricated and debugged, 10 of which are presently operational in the field. The boards are populated with the factory-programmed ROM SPIs. Four are in use in PVTs at Information Sciences Institute (ISI), SRI International, and Defense Communications Engineering Center (DCEC). Six are in use in PVTs at Lincoln.

Two of the LPC vocoders at Lincoln and two vocoders at ISI and SRI were successfully used in a four-party conference over the wideband network in a rehearsal for the 3 June packet speech demonstration. During the actual demonstration on 3 June, the SRI site used the CHI-V LPC implementation thereby demonstrating the compatibility of the Lincoln and Culler-Harrison, Inc. (CHI) LPC vocoders.

The PVT telephone handsets use a passive microphone element produced by Telephonics. This element, which performs adequately for PCM and embedded CVSD (ECVSD) voice digitization, has been found to significantly degrade LPC vocoder performance. Currently a microphone element made by Roanwell, which has been found to be satisfactory for use with LPC, is being used to replace the Telephonics units in PVT handsets that communicate with the LPC vocoder. Unfortunately, the Roanwell element is no longer commercially available and Lincoln has only a small number in-house. Recently, we have been able to purchase samples of a new passive microphone element produced by Roanwell. Informal tests using this element indicate satisfactory performance with the

LPC vocoder. If the element passes more rigorous tests it will be used as the new standard PVT microphone replacing both the Telephonics and older Rosmwell units.

Detailed hardware and software designs have been prepared for a generalized stand-alone version of the compact LPC to include self-contained audio and an RS-232 interface for connection to host computers. These designs are based partially on collaborative effort with ISI, particularly on the specification of the protocol on the serial RS-232 link. In addition, comments received from other members of the DARPA packet speech community on an earlier functional design document have been factored into the detailed design. We expect the initial hardware and software development and checkout to be completed during the first quarter of FY 83.

II. COMPACT EMBEDDED VOCODER

A compact multirate LPC vocoder system has been developed which (a) operates as an embedded speech coder in the 800- to 4800-bps range, and (b) can be implemented in hardware using, without change, the NEC SPI analyzer, synthesizer, and pitch detector chips developed for the compact LPC (discussed in Sec. I). The mechanism by which embedded coding has been effected has been to develop an interface between the LPC algorithm and the external communication system, of which the LPC analyzer and synthesizer are totally ignorant. The interface is therefore not restricted to the NEC SPI-based LPC realization, but could operate with any equivalent LPC realization. Both five-rate and three-rate systems have been developed. The vocoders were first implemented as real-time systems on the Lincoln Digital Signal Processor (LDSP). Architectures were developed for hardware realizations using the NEC SPIs and an INTEL 8085-based microcomputer for the interface between the vocoder chips and the PVT. Microcode was developed for the 8085, and it was shown that a single 8085 would support either the five- or three-rate system in real time.

The five-rate system will be described first, along with the basic analyzer and synthesizer interfaces. The system operates at the five rates of 889, 1244, 2489, 3556, and 4622 bps. This particular set of bit rates is explained by the fact that the embedded vocoder is designed to communicate with a packet network in which parcels must be multiples of 8-bit bytes. This design assumes an analyzer which updates LPC parameters every 11.25 ms. As noted above, the mechanism by which embedded coding is effected essentially involves an interface between the LPC algorithm and the external world of which the LPC analyser and synthesizer are totally ignorant. The different rates are obtained by application and extension of frame-fill concepts[1,2] to conform with the constraints of embedded coding.

A description of the analyser interface follows. As shown in Fig. 1, raw parameters received from the LPC analyzer are coded and buffered until four sets of parameters (A to D) representing 45 ms have been accumulated.

(The pitch period is coded logarithmically as in the DoD standard algorithm,[3] and the energy and K-parameters are coded using tables from the Lincoln autocorrelation LPC (Ref. 4), which has been utilized in most of the DARPA packet speech experiments. Using the four sets of coded parameters, there are now three tasks to be performed. First, sets of truncated pointers from frames A, C, and E are used to decide how best to reconstruct the parameters of frame C in the event that the synthesizer interface receives only priority I information. The priority I parcel is then formed which contains the minimum information needed, as seen in Table I. Second, the sets of full pointers from frames A, C, and E are used to make the frame-fill decision for frame C. This information plus the LSBs of the truncated pointers become the priority II parcel (see Table II). Third, the sets of pointers from frames C, D, and E are used to determine the frame-fill strategy for frame D in the event that the parameters of frame C reach the synthesizer interface. The contents of the priority III parcel are shown in Table III. (Note that the LSBs of the pitch period pointers for frames B and D are included.) Priority IV and V parcels containing the parameters for frames B and D (shown in Tables IV and V) are then formed. All parcels are now ready for transmission to the protocol processor.

The synthesizer interface receives from the protocol processor either all (I through V) or a stripped subset (I through ?) of the transmitted parcels and the indication of which subset has survived transmission (V,IV,...,I). If only priority I has survived, reconstruction of frames B, C, and D is as follows (refer to Fig. 2). The coded parameters for frame C are formed either by copying from frame A or E or interpolating between them as dictated by the frame-fill strategy included in the priority I information. The pitch pointers of A and C are copied into B and D, respectively.

The energy and K pointers of B and D are formed by interpolating the adjacent frames as shown. The four frames of parameters are then decoded using a special set of decoding tables with coarser quantization. Reconstruction of the four frames is the same if priority II information has

<table>
<tr><th colspan="8">TABLE I

PRIORITY I PARCEL (5 Bytes)</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="2">E (A)
4 MSBs</td><td colspan="6">Pitch (A)</td></tr>
<tr><td colspan="2">K2 (A)
4 MSBs</td><td colspan="4">K1 (A)
4 MSBs</td><td colspan="2"></td></tr>
<tr><td colspan="2">K4 (A)
3 MSBs</td><td colspan="4">K3 (A)</td><td colspan="2"></td></tr>
<tr><td colspan="2">K7 (A)
2 MSBs</td><td colspan="2">K6 (A)
2 MSBs</td><td colspan="3">K5 (A)
3 MSBs</td><td></td></tr>
<tr><td colspan="3">Frame-Fill
Strategy (C)</td><td>V/UV
(C)</td><td>K10 (A)
1 MSB</td><td>K9 (A)
1 MSB</td><td colspan="2">K8 (A)
2 MSBs</td></tr>
</table>

<table>
<tr><th colspan="8">TABLE II

PRIORITY II PARCEL (2 Bytes)</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>K6 (A)
2 LSBs</td><td>K5 (A)
1 LSB</td><td>K4 (A)
1 LSB</td><td colspan="2">K2 (A)
2 LSBs</td><td colspan="2">K1 (A)
2 LSBs</td><td>E (A)
1 LSB</td></tr>
<tr><td colspan="3">Frame-Fill
Strategy (C)</td><td>K10 (A)
1 LSB</td><td>K9 (A)
1 LSB</td><td>K8 (A)
1 LSB</td><td>K7 (A)
1 LSB</td><td></td></tr>
</table>

TABLE III

PRIORITY III PARCEL (7 Bytes)

<table>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="2">---E (C)</td><td colspan="6">Pitch (C)</td></tr>
<tr><td colspan="5">-------K1 (C)</td><td colspan="3">←-------</td></tr>
<tr><td>--K3 (C)</td><td colspan="6">K2 (C)</td><td>←---</td></tr>
<tr><td>--K5 (C)</td><td colspan="4">K4 (C)</td><td colspan="3">←-------</td></tr>
<tr><td>--K7 (C)</td><td colspan="4">K6 (C)</td><td colspan="3">←-------</td></tr>
<tr><td>--K10 (C)</td><td colspan="2">K9 (C)</td><td colspan="3">K8 (C)</td><td colspan="2">←-----</td></tr>
<tr><td>X</td><td>Pitch (D)
1 LSB</td><td>Pitch (B)
1 LSB</td><td colspan="3">Frame-Fill
Strategy (D)</td><td>V/UV
(D)</td><td>←---</td></tr>
</table>

TABLE IV

PRIORITY IV PARCEL (6 Bytes)

<table>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="3">-----E (B)</td><td colspan="5">Pitch (B)
5 MSBs</td></tr>
<tr><td colspan="6">K1 (B)</td><td colspan="2">←-----</td></tr>
<tr><td colspan="2">-----K3 (B)</td><td colspan="6">K2 (B)</td></tr>
<tr><td colspan="2">-----K5 (B)</td><td colspan="4">K4 (B)</td><td colspan="2">←-----</td></tr>
<tr><td colspan="2">-----K7 (B)</td><td colspan="4">K6 (B)</td><td colspan="2">←-----</td></tr>
<tr><td colspan="2">K10 (B)</td><td colspan="2">K9 (B)</td><td colspan="3">K8 (B)</td><td>←---</td></tr>
</table>

TABLE V

PRIORITY V PARCEL (6 Bytes)

E (D)			Pitch (D) 5 MSBs				
K1 (D)							
K3 (D)		K2 (D)					
K5 (D)		K4 (D)					
K7 (D)		K6 (D)					
K10 (D)		K9 (D)		K8 (D)			

been received except that the final decoding of the parameters is done using the regular set of decoding tables. Figure 3 shows frame reconstruction when priority III parameters are received. Frame D is formed as dictated by the frame-fill strategy for D, and frame B is constructed as described above. If priority IV information has been received, only frame D has to be filled before decoding the parameters, as shown in Fig. 4. Priority V information necessitates only decoding the parameters. Decoded sets of parameters are then buffered and sent to the synthesizer every 11.25 ms. The decoding tables for the energy and K-parameters are double length to accommodate the interpolation of pointers.

This five-rate embedded LPC vocoder was first implemented as a real-time system on the LDSP. Diagnostic Rhyme Test (DRT) scores of vocoder performance were obtained for the following rates:

Rate (bps)	DRT Score
889	84.4
1244	87.2
2489	90.6
4622	91.1

Scores were averaged over three speakers.

Our ultimate goal was to assess the feasibility of using an INTEL 8085-based 8-bit microcomputer to support the interface between the LPC vocoder in the NEC chips and the Packet Voice Terminal. Figure 5 depicts the suggested architecture for a five-rate embedded LPC-10. Notice that an additional analyzer chip would be needed for the computational capacity to provide a new set of reflection coefficients every 11.25 ms. The vocoder interface was completely programmed for the 8085 and verified with canned data on the emulator of the Hewlett-Packard development system. Despite initial fears that a single 8085 would not suffice, it appears that if a 200-ns clock were used real time would not be exceeded. Memory requirements for this interface are:

ROM = 6K

RAM = 1K

The five-rate embedded LPC-10 interface was then stripped to represent only three rates: 889, 1244, and 2311 bps. This interface does not require an additional NEC analyzer chip because frames are 22.5 ms in length instead of 11.25 ms. The contents of the three priority parcels may be seen in Tables VI through VIII. Memory requirements are:

ROM = 4K

RAM = 1K

Both the five-rate and three-rate LPC analyzer interfaces include silence detection. If the residual energy has not exceeded a threshold during the last 300 ms, this information is conveyed to the protocol processor via the 8-bit header. The energy threshold is adaptive in order to accommodate differences in handsets and/or ambient noise. The algorithm for adaptation is deliberately simplistic. Every N seconds a new minimum energy is established. All frames which are either voiced or exceed the energy threshold by M decibels are declared speech. N has been arbitrarily set to 2 s, and M is equal to 4.5 dB. Since the quantized energies of the coding table are spaced a fixed decibel apart, the algorithm deals solely with pointers to the coded values, eliminating any need for 16-bit arithmetic.

<table>
<tr><th colspan="8">TABLE VI

PRIORITY I PARCEL (5 Bytes)</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="2">E (A)
4 MSBs</td><td colspan="6">Pitch (A)</td></tr>
<tr><td colspan="2">K2 (A)
4 MSBs</td><td colspan="4">K1 (A)
4 MSBs</td><td colspan="2"></td></tr>
<tr><td colspan="2">K4 (A)
3 MSBs</td><td colspan="4">K3 (A)</td><td colspan="2"></td></tr>
<tr><td colspan="2">K7 (A)
2 MSBs</td><td colspan="2">K6 (A)
2 MSBs</td><td colspan="3">K5 (A)
3 MSBs</td><td></td></tr>
<tr><td colspan="3">Frame-Fill
Strategy (B)</td><td>V/UV
(B)</td><td>K10 (A)
1 MSB</td><td>K9 (A)
1 MSB</td><td colspan="2">K8 (A)
2 MSBs</td></tr>
</table>

<table>
<tr><th colspan="8">TABLE VII

PRIORITY II PARCEL (2 Bytes)</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>K6 (A)
1 LSB</td><td>K5 (A)
1 LSB</td><td>K4 (A)
1 LSB</td><td colspan="2">K2 (A)
2 LSBs</td><td colspan="2">K1 (A)
2 LSBs</td><td>E (A)
1 LSB</td></tr>
<tr><td></td><td colspan="3">Frame-Fill
Strategy (B)</td><td>K10 (A)
1 LSB</td><td>K9 (A)
1 LSB</td><td>K8 (A)
1 LSB</td><td>K7 (A)
1 LSB</td></tr>
</table>

<table>
<tr><th colspan="8">TABLE VIII

PRIORITY III PARCEL (6 Bytes)</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="2">----E (B)</td><td colspan="6">Pitch (B)</td></tr>
<tr><td colspan="5">--------K1 (B)</td><td colspan="3">←--------</td></tr>
<tr><td>--K3 (B)</td><td colspan="6">K2 (B)</td><td>←---</td></tr>
<tr><td>--K5 (B)</td><td colspan="4">K4 (B)</td><td colspan="3">←--------</td></tr>
<tr><td colspan="2">---K7 (B)</td><td colspan="3">K6 (B)</td><td colspan="3">←-------</td></tr>
<tr><td colspan="2">K10 (B)</td><td colspan="2">K9 (B)</td><td colspan="3">K8 (B)</td><td>←--</td></tr>
</table>

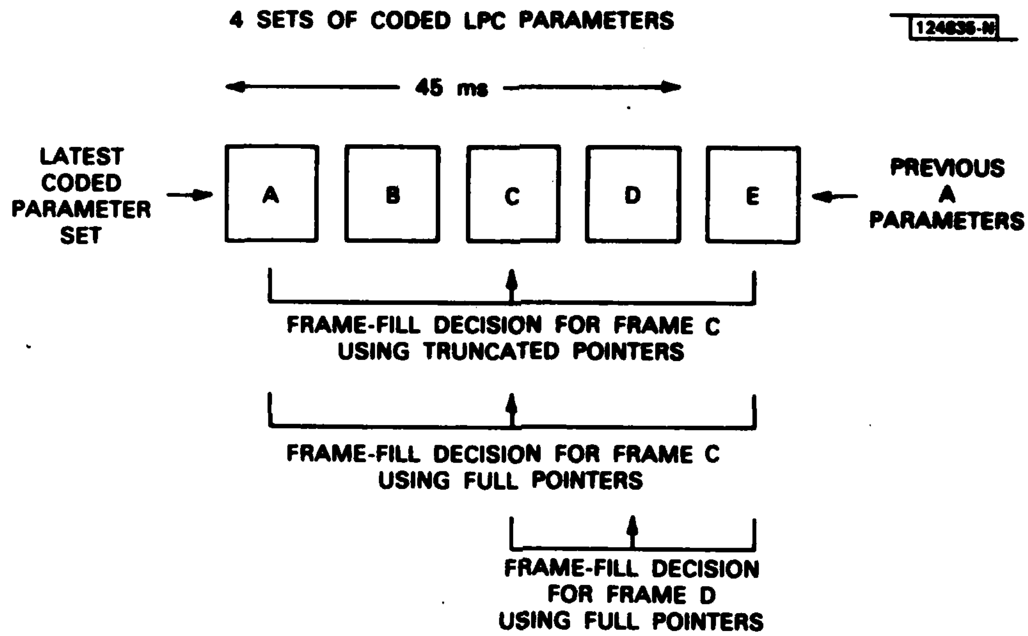

Fig. 1. Analyzer interface.

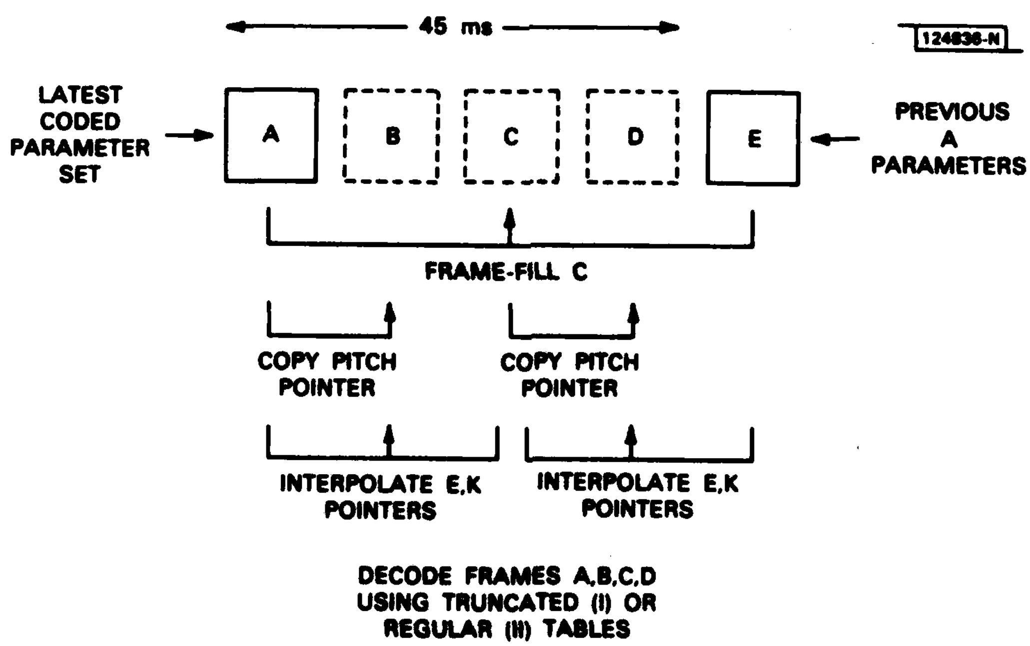

Fig. 2. Priority I and II synthesis.

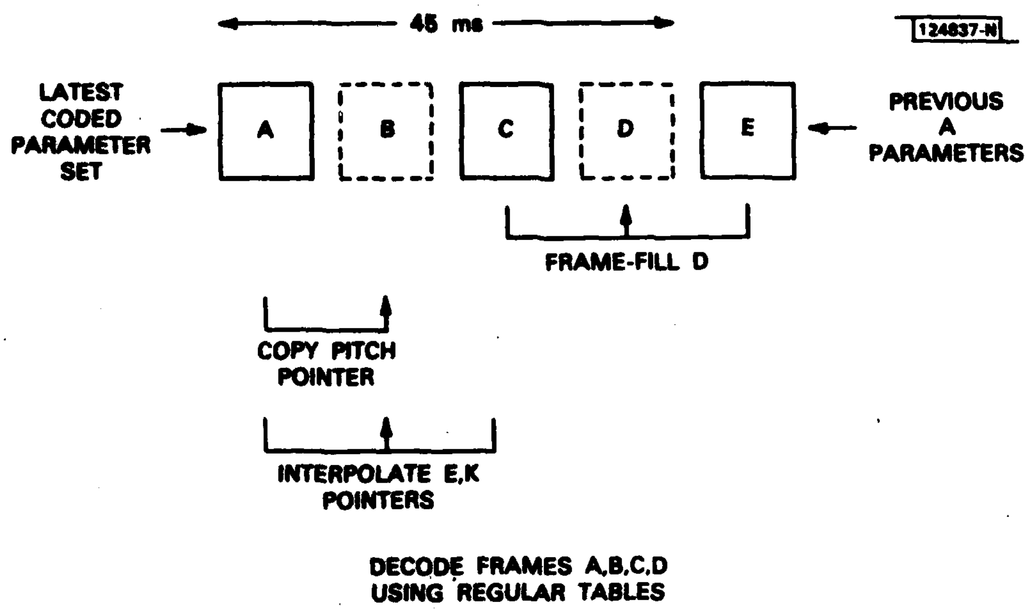

Fig. 3. Priority III synthesis.

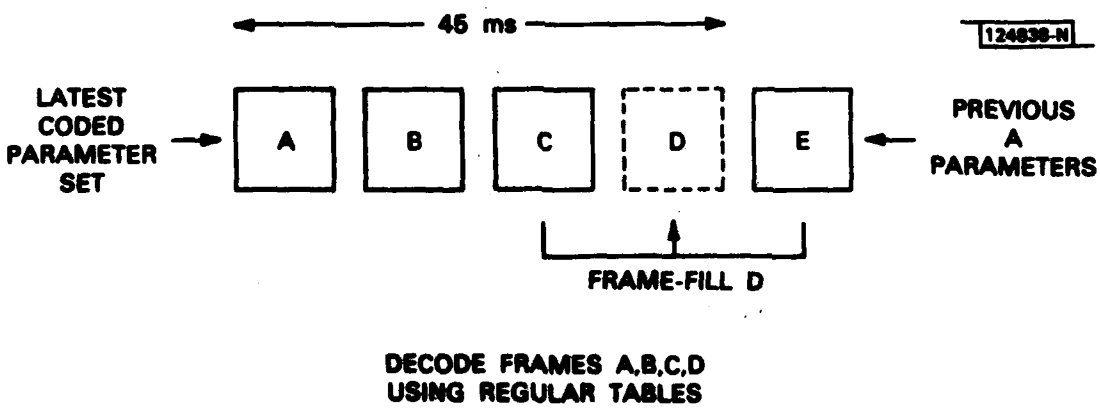

Fig. 4. Priority IV synthesis.

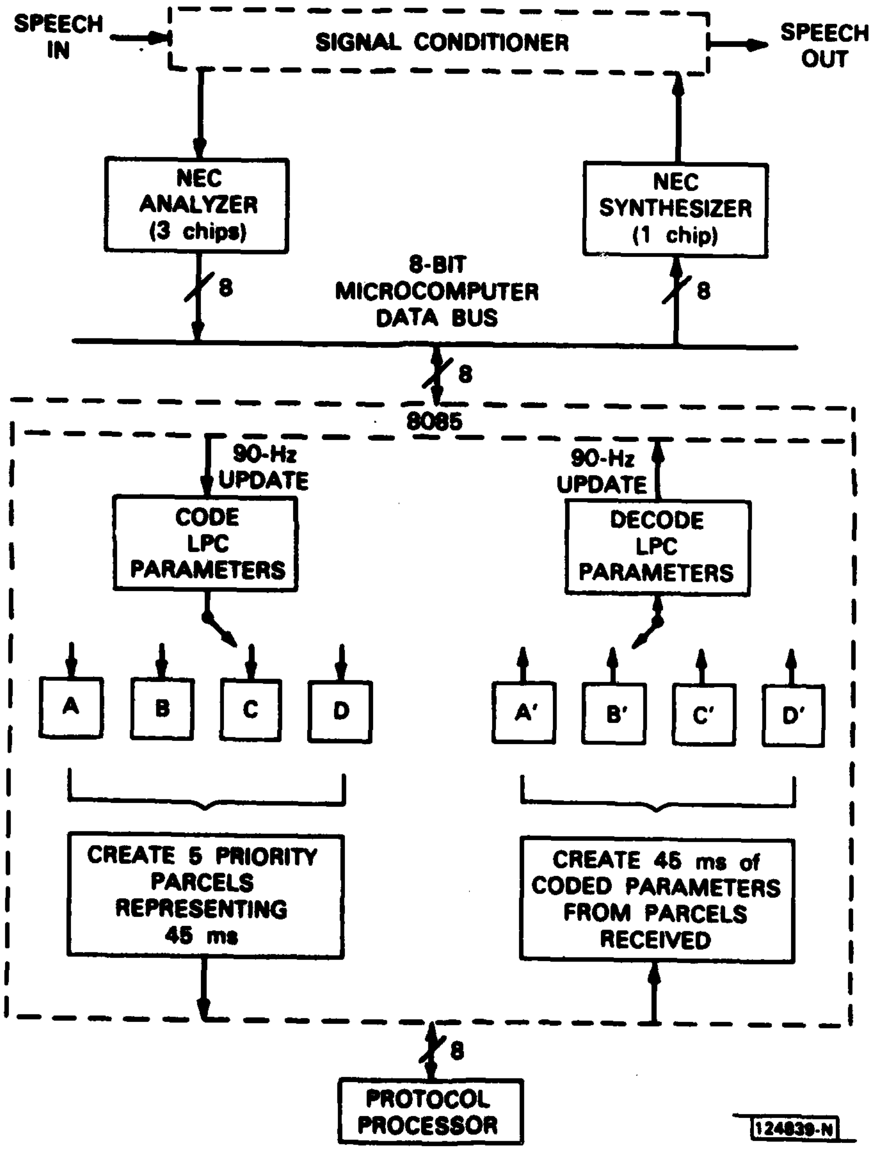

Fig. 5. Embedded LPC architecture.

III. PACKET VOICE TERMINAL

A. PVT AND LEXNET HARDWARE STATUS

Two additional LEXNET Concentrator Interface (LCI) units have been constructed to allow the installation of networks at new sites while maintaining the capability of performing two-network tests at Lincoln.

LEXNETs have been installed at SRI and DCEC, each consisting of an LCI and one PVT. They were initially installed to support the 3 June demonstration, but now remain in place for use in general testing of the satellite channel. Another single-terminal LEXNET has been returned to Bolt, Beranek and Newman (BBN) for use in testing of the Voice Funnel.

Several hardware upgrades have been installed in the LEXNET equipment over the last quarter. The most significant of these was a change to increase the data rate on the channel from the LCI to the concentrator from 375 to 600 kbps. This change is possible because of improved software for handling the SIO chip on the UMC-Z80. Further experiments were conducted to see if this rate could be increased further. It was found that this link would start to fail at about 625 kbps, but would work reliably at lower rates; the current LCI cards are set to run at 600 kbps. The remaining changes involved minor wiring additions to allow the installation of the timer card in the PVT.

B. PVT TECHNOLOGY TRANSFER

A draft specification for the PVT has been written and is undergoing internal review. The specification asks for terminals which are electrically identical to the current terminals. We are requiring that cards be usable in the current terminals for testing and spares, but we give the contractor the option of translating the design to a printed circuit card. The Request for Quote (RFQ) will ask for prices at quantities of 10, 20, and 50 terminals.

C. PVT SOFTWARE STATUS

The PVT conferencing feature is now operational. The PVTs handle three vocoders: PCM, 2400-bps LPC (NEC SPI-based vocoder), and embedded Continuously Variable Slope Delta Modulation (ECVSD). The ECVSD can run at rates at 16, 32, 48, and 64 kbps. The rate can be changed at any time, even in mid-talk spurt.

A distributed "floor controller" is operational. The algorithm is included in the PVT software and resolves conflicts when more than one participant tries to talk at the same time. The PVT will not transmit speech if its user does not "have the floor." Conflicts are currently resolved on a simple priority basis but the floor controller is designed so that other algorithms may be easily substituted.

A time out and retransmission mechanism, for enhanced robustness in call setup, has been incorporated into the PVT software. The PVT monitors each control message sent and retransmits a message when the appropriate response is not received. After ten consecutive failures the PVT will automatically close down the connection. This feature should ensure that conference and point-to-point (PTP) connections are reliably established even during times of high packet loss. This algorithm has been debugged locally between PVTs and between PVTs and the Access Controller. It must still be tested over the satellite.

The Access Controller (AC) has been operating successfully for a variety of conference scenarios. As a system resource, the AC software is running in a PVT at a "well known address" on a LEXNET at Lincoln. Currently, the AC knows about three possible conferences (one for each possible vocoder type). It currently accepts all callers who know the conference name and password and who can use the correct vocoder. Other types of conferences may be added easily.

The AC has the ability to operate with the new PVT software. It facilitates the retransmission mechanism by acknowledging all received messages and it handles the duplicate messages it occasionally receives because of retransmissions.

Conferences may be set up in three ways. All the participants can dial the AC and ask to join the conference (referred to as "Meet Me"), a conference originator can invite in others that he wishes to have in his conference, or the conference originator can dial the Voice-Controlled Operator (VCOP) and request that a conference be set up.

In the "Meet Me" style of conferencing each participant dials the AC. The AC provides information about the others in the conference and PVT software executes the necessary protocol steps to set up connections to all the other conferees. This was the first method of conferencing implemented and was demonstrated at the 3 June packet speech meeting.

A new dial up method of starting a conference is now working. The conference originator after entering the conference himself, dials in turn the PVT addresses of the other participants. His PVT sends a special "Please Join a Conference" to the other PVTs. A PVT receiving this message rings its phone. When the phone is answered the PVT initiates the protocol exchange with the AC which lets the individual join the conference.

Based on the code written to implement "Please Join a Conference," the first stage of the PVT software necessary to support the Voice-Controlled Operator (VCOP) is written and largely debugged. A conference originator calls the VCOP PVT. When the protocol exchange is finished, VCOP asks the originator a series of questions to obtain necessary conference parameters. VCOP then requests the originator to hang up his phone. VCOP passes to its PVT the parameters it has received from the originator. Its PVT then issues "Please Join a Conference" messages to the participants' PVTs. Section VI discusses VCOP status in more detail.

Other features have been added to make the PVTs more robust. Closing down a conference can take more than 20 s if the PVT cannot get its disconnect message through to some other participant (the disconnect is retransmitted ten times at 2-s intervals). If a caller picks up the phone during this period he will hear a fast busy signal and his dialing will be ignored. As soon as the shutdown protocol has completed, the caller will get a dial tone and can place his call.

Software in both the PVTs and the AC has been upgraded to correctly handle the case where a PVT crashes and then tries to reenter the conference. A PVT can hang up and later rejoin a conference at will. During a conference which has been set up in any of the three ways listed above, any participant can "Ask In" another PVT by dialing its PVT address.

D. MEASUREMENT HOST DEVELOPMENT

A measurement host capability has been added to the PVT, in order to augment our facilities for carrying out performance measurements on the wideband network. The most important difference between this measurement host and previous PVT-based traffic emulators is the capability for precise timing measurements on cross-net packet transmissions. This capability is provided by means of a new timer card which plugs into the vocoder slot on the PVT. The timer card, in turn, connects to a precise global timing clock which has been provided to us by ISI. The clock obtains precise global time from a radio signal (WWVB) maintained by the National Bureau of Standards. Similar WWVB clocks are installed at the ISI site. In addition to the timer card, the measurement host also includes new software in the protocol processor which provides a flexible capability for packet generation, delay histogram collection, and configuration of experiments.

The user can configure experiments based on the following set of capabilities:

(1) Traffic Models

Deterministic - The host generates fixed-length packets of L bytes at a steady rate R per second. R and L are selectable at set-up time.

Poisson - Same as deterministic except R is now taken to be a mean rate in emulating a Poisson process.

Talker Activity Model - The host emulates a statistical model of talker activity for each of N talkers. Every J frames (1 frame = 20 ms) it examines the activity of each simulated talker over the previous J frames and sends a packet of L bytes if activity occurred. The user can select N, J, and L, and certain parameters of the activity model.

(2) Packet Protocol

Either the Internet datagram Protocol (IP) or Stream (ST) Protocol can be selected.

(3) Output

The receiving host keeps a histogram of the absolute delay or delay dispersion depending on whether both hosts have access to a common clock (see below). At set-up time the user can select the histogram bin size.

A typical experimental configuration is illustrated in Fig. 6. The entire experiment can be controlled from one of the hosts. In Fig. 6 the user communicates directly with A via terminal and can ask A to send control messages to B as well.

The label "WWVB" refers to the WWVB radio signal (maintained by the NBS) which can be used to synchronize clocks. The measurement host is designed to run with the SPECTRACOM Model 8170 WWVB synchronized clock. If these clocks are available to both hosts, one obtains one-way, absolute delay measurements. Otherwise, one or both hosts must use an internal clock, and they measure one-way delay dispersion and round-trip absolute delay.

The procedure in a typical experiment is as follows.

(1) The user at A tells B to be a receiver for packets from A. At this point, B zeroes its counts and waits for packets from A.

(2) The user tells A to transmit to B, selecting one of the options described before. A then transmits to B, selecting one of the options described before. A then transmits time-stamped packets to B.

(3) The user tells A to stop transmitting.

(4) The user then retrieves results. Host B is told to send its results from display on A's terminal. The histogram appears in tabular form. Host A can be asked to display the number of packets sent for comparison.

The timing and traffic emulation functions of the measurement host are performed by a microprocessor-based timing card which occupies the speech processor slot of the PVT. The main measurement host program, which interacts with the user and constructs the needed packets and headers, replaces the NVP processor program of an ordinary PVT.

The hardware is complete and preliminary experiments have been run using IP packets. Work continues on adding an ST packet capability to the card.

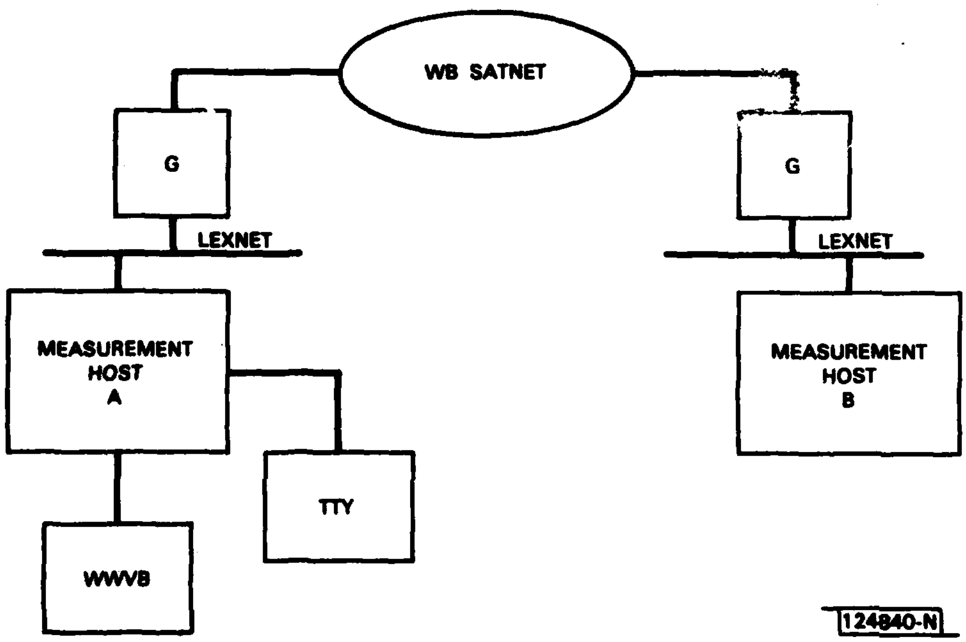

Fig. 6. Measurement host experiment configuration.

IV. MINICONCENTRATOR

A. HARDWARE STATUS

Throughout this semiannual reporting period miniconcentrator gateways have been running at ISI, SRI, DCEC, and Lincoln Laboratory. The gateways at SRI and DCEC were augmented with hardware to support LEXNETs so that a richer packet voice environment would be available for demonstrations and experiments. All the gateways now in service have three UMC-Z80 interface processors connected to a central PDP-11/44 miniconcentrator.

Our experience with the reliability of the miniconcentrator hardware has been very encouraging. In the past six months we have experienced no hard failures in the eight PDP-11/44 machines with which we have been working. We have experienced an intermittent problem with one of the machines that has been in use in our laboratory as a checkout facility. Over the same period we have experienced four hard failures in 35 UMC-Z80 processor and memory boards and a similar number of intermittent problems. One of these hard failures caused an outage of several days for one of the gateway machines, but otherwise there were no gateway hardware malfunctions that caused interruptions in network experiments.

B. GATEWAY SOFTWARE

1. *Gateway Program*

Several major extensions to the PDP-11 gateway program were achieved during this time period. These extensions are in the areas of speech conferencing, robustness, measurement and instrumentation experiments, and more extensive external control of the gateway program.

Foremost among the extensions was multisite speech conferencing between participants on local networks communicating via gateways and the PSAT. In order to establish a conference the gateways use ST protocol to communicate with each other the control messages that were originated by LEXNET PVTs and

PRNET SIUs. Then the "bit maps" in the speech messages are used to control the forwarding of speech packets generated by the participants. This capability culminated in a demonstration at the packet speech meeting at Lincoln on 3 June in which the participants in the conference were at a PVT on a LEXNET at Lincoln, at a PVT on a LEXNET at ISI, and at an SIU on a PRNET at SRI. The gateways at each site broadcast the LPC speech to each other using stream allocations and group addressing.

Several prior achievements were instrumental in the success of the demonstration. An ad hoc protocol was developed by means of which each gateway communicates with the others both to tell them that it is on the air and to request them to join a PSAT group that it has created. (PSAT groups are a mechanism provided by PSATs in which several hosts who belong to the group receive a message sent to the group address; these groups are then used as the broadcast addresses for the transmission of the speech from one participant to all the others.) Additionally, a network module for the gateway to handle PRNET communication was generated and several gateways were installed on the SRI PDP-11/44 computer to handle various combinations of PSAT, LEXNET, and PRNET networks.

To provide greater robustness in establishing connections and conferences a time-out/retransmission mechanism has been developed in the gateway. Using a state table the gateway retransmits ST control messages a number of times until they are correctly acknowledged or until a fixed number of transmissions have been performed. If success is not achieved then the gateway proceeds to recover from the missing control messages.

In order to assist in the carrying out of measurement and instrumentation experiments the handling of new-style IP source-routing (a revised standard in the DARPA Internet community) was incorporated in the gateway, replacing an older obsolete source-route option. The new source-routing is useful for experimentation with and measurement of delays involved in using different network paths, and in particular allows us to source-route through IP gateways (developed by BBN and others) located elsewhere in the DARPA Internet. Additionally, the gateway now monitors its internal memory resources in histograms during its operation. This enables one to determine the extent of spare capacity for various experiments with the goal of balancing resources.

A protocol was developed and implemented between the gateway program running in the PDP-11 and the network input-output programs running in the UMC interface processors. This protocol, together with matching keyboard gateway commands, permits one to examine and modify UMC memory while the program is running. This provides for convenient access to counts that the UMC programs maintain during their operation. Previously, the UMC programs had to be halted in order to perform such operations.

In certain error cases the gateway now generates Internet Control Message Protocol error messages to the source of the incorrect message.

The Exploratory Data Network (EDN) was added to the list of networks that the gateway supports.

For external gateway control a number of new commands were added. Included are commands to specify the PSATs that are to be involved in group addresses, to create and delete streams, to clear out certain information that the gateway is maintaining, and to dynamically control the gateway mode bits. For logging and monitoring purposes the date-time is supplied in critical places in terminal output produced by the gateway. Finally, for maintenance purposes each gateway has a version number and unique serial number associated with it.

2. Support Software

The downline communication program TOEPOS was installed on 2 host computers at SRI running under V7 UNIX and communicating with the object computer via a MICOM terminal multiplexer. This marked the fourth distinct operating environment in which the program has been installed.

TOEPOS was extended to accommodate downline computers which do not run the EPOS operating system. If such use is specified when TOEPOS is invoked, then the functions relating to multiplex TTY line selection are suppressed, yielding type-through communication free of TTY line identifications. Using this mode, TOEPOS has subsumed the functions of the outdated TOLEXNET program for communicating with and downloading PVTs. Additionally, TOEPOS is now being used for running PVT Measurement Host experiments and controlling the generation of subsequent plots of the results.

A command was added to TOEPOS by means of which a user can request TOEPOS to supply its terminal output to an additional user-specified terminal line. Such output is in addition to that supplied to the terminal on which TOEPOS is running and to the script file. This provides a "watch" capability in which a remote site can watch the actions at the terminal running TOEPOS. This capability was used in the 3 June demonstration to provide instantaneous awareness of the state of the gateways to observers at Lincoln and the site at which the gateway was running.

V. WIDEBAND NETWORK EXPERIMENTS AND EXPERIMENT COORDINATION

The major emphasis in experiment planning and coordination in the April-June time frame was preparation for a wideband packet speech demonstration to be held at Lincoln on 3 June. The general goals of the demonstration had been defined at the wideband meeting at DARPA in March, accompanied by extensive discussion of the tasks and integration activities that would have to be completed throughout the network in order to achieve these goals. DARPA inaugurated a system of weekly reporting of the status of all of these activities, with Lincoln acting as the coordination point for all the sites. The weekly reports contained statements of the goals that had to be achieved with respect to WB SATNET operation and integration of various equipment at the sites, and included summaries of all intermediate milestones achieved to date in each area.

The packet speech demonstration itself, which was conducted during the 3 June DARPA packet speech meeting at Lincoln, was an excellent success. A broad set of capabilities for packet speech communication over WB SATNET, with internetwork connections to LEXNETs and a packet radio net (PRNET), was shown by means of a sequence of calls among the Lincoln (LL), ISI, and SRI sites. The demonstration followed a period of intense effort and cooperation among all the participants in the Wideband Program, which culminated in a series of dry runs which were very useful in final integration of all the subsystems involved.

The equipment configurations that were available at the four network sites at the time of the demonstration are summarized in Fig. 7. The succession of demonstration calls, to be described below, will be described in the context of Fig. 7.

As indicated, each of the four sites had its earth station, ESI, PSAT, and various hosts. SRI had a miniconcentrator gateway; a LEXNET with one Packet Voice Terminal (PVT) equipped with a Lincoln single-card 2.4-kbps LPC vocoder in addition to the built-in 64-kbps mu-law PCM coder; and a Packet Radio Net with two PRNET packet voice terminals, each consisting of an LSI-11/23 Speech Interface Unit (SIU) and a CHI-V vocoder running the

2.4-kbps LPC-10 algorithm. One of these voice terminals was mounted in a van, and participated in the mobile PRNET packet voice demonstration below. ISI had a Voice Funnel configured as a gateway (which was not involved in the demonstration); a miniconcentrator; a LEXNET with two PVTs; and a PDP-11/45 speech host with packet voice and video capability (which were also not involved in the demonstration). One PVT was equipped with an ISI-built Switched Telephone Network Interface (STNI) card which enables a 64-kbps PCM packet voice user on the network to access the public telephone system, as described below, and the other was equipped with the single-card LPC vocoder. Lincoln had a Funnel (not used in the demonstration); a miniconcentrator and LEXNET, including multiple PVTs and the Conferencing Access Controller described below; and a Packet/Circuit Interface (PCI) and Telephone Office Emulator, or TOE (developed under DCA sponsorship, and not involved in the demonstration). Two of the PVTs were equipped with 16- to 64-kbps Embedded CVSD (ECVSD) vocoder cards, and two had LPC vocoder cards. The DCEC site, equipped as shown, was not included in the 3 June demonstration; DCEC was the site of another demonstration in early October focusing on the PCIs and TOEs.

As indicated in the generic packet voice call demonstration setup depicted in Fig. 8, audio pickups were attached to various handsets at Lincoln, so that the reconstituted voice signals from the various participants could be played through the conference room audio system for the benefit of the audience. Figure 8 also illustrates the point that each call required two classes of packets to be transmitted between the terminals: (a) control packets to set up and take down the call; and (b) speech packets to transport the digitized voice.

The demonstration sequence of calls which were completed is summarized in Table IX. The demonstration sequence began with four local calls on the LEXNET at Lincoln, as follows: a point-to-point call using 64-kbps PCM; a comparison of voice quality at the various rates in a point-to-point ECVSD call; a point-to-point 2.4-kbps LPC call; and a 4-party conference using 64-kbps PCM. Two point-to-point calls over the satellite net were then completed: a 64-kbps call between a PVT at Lincoln and one at ISI, and a 64-kbps call from Lincoln to the local Los Angeles weather forecast number by

TABLE IX

SUMMARY OF PACKET SPEECH DEMONSTRATION SEQUENCE

3 JUNE 1982

LOCAL CALLS ON LINCOLN LEXNET

Point-to-Point

1. PCM (64 kbps)
2. ECVSD (16 to 64 kbps)
3. LPC (2.4 kbps) Using Single-Card LPC in PVT

Conference

4. PCM 4-Party Conference

CALLS OVER WIDEBAND SATNET

Point-to-Point

5. PCM:LL - ISI
6. PCM:LL - Switched Telephone Network Interface at ISI (call to local Los Angeles weather)

Conference

7. 3-Site, 4-Party LPC Conference
   LL (2 parties on LEXNET)
   ISI (on LEXNET)
   SRI (on PRNET unit in speech lab, using CHI-V LPC and SIU)

Point-to-Point (mobile)

8. LPC:SRI (mobile PRNET unit) - LL (LEXNET PVT)
   Mobile Van Run Demonstrating PRNET Speech and Alternate Routing

way of the STNI card at ISI. A 3-site, 4-party LPC conference over the WB SATNET was then set up, involving two PVTs at Lincoln; a PVT at ISI; and a PRNET packet voice terminal in the speech laboratory at ISI. Finally, an extended mobile call was set up between a PVT at Lincoln and the moving packet radio-equipped van on local streets and the Bayshore Freeway near SRI, using 2.4-kbps LPC via a PRNET packet voice terminal in the van. During the latter run, signals from the van were sent to the PRNET base station at SRI via single and multiple-hop routes, as determined by obstructions in the line-of-sight paths, and were relayed from the base station to Lincoln via the WB SATNET.

An important consideration throughout the preparations for the 3 June demonstration was Western Union activity in upgrading the remote earth station monitoring and control equipment and resolving a few known problems with the earth station equipment at the various sites. At the time of initial installation a temporary expedient had been taken for remote monitoring at each of the four sites, because the higher-quality equipment normally used by Western Union was not available on the necessary time scale. It was agreed with Western Union in March that a team would visit each site in turn to complete the installation, with scheduling and coordination carefully handled to avoid conflicts with testing of speech equipment for the 3 June demonstration. This effort was largely successful; the new equipment was put in place at all sites, and it was fully activated and checked out at all sites except ISI. Various problems and delays had impeded completion of that final step until a short time before the demonstration, and Western Union was asked to defer further work to give maximum time for testing and preparation for the demonstration.

By agreement with Western Union, it has been determined that Lincoln Laboratory will act as the transmitted frequency and power reference station for the Wideband SATNET. The requirement for such a reference is based on the fact that these two parameters must be identical at all stations in the network, within rather narrow limits, to accommodate the burst-to-burst AGC and AFC windows of the ESIs and the FCC power allocations assigned to Western Union. Initial alignment of these parameters is difficult, involving transportation of calibration equipment to each site, while tracking and correcting long-term drifts pose additional problems.

The chosen solution for these problems is to install an accurate spectrum analyzer with an internal frequency reference at one site, and to periodically verify that each of the other stations transmits at the same frequency and power, as measured at the reference site. This implies that signals arriving at the satellite will be identical, and that the receiver at any given site (although it may have an uncompensated and unimportant internal offset) will receive all the signals without burst-to-burst variation. Performing these periodic checks will require cooperation between BBN and Lincoln, in that the station under test must be the only one operating specific repetitive data patterns that can be conveniently tracked by the spectrum analyzer.

A major objective of current wideband experimental activity is to achieve increased system throughput in order to support multiplexing of large numbers of users. The network has typically been operated at a burst rate of 772 kbps up to the present, thus keeping the bit error rate sufficiently low to remove it from consideration in isolating and correcting bugs in the various gateways and speech equipment. More recently, with steadily increasing confidence in this equipment, various combinations of higher bit rates with and without coding are being tested for extended periods. Concurrently, experiments are being planned and performed to compare the multiplexing performance of the system with the predictions embodied in W-Note-33, "Wideband Network Throughput Calculations," distributed and discussed at the March Wideband Meeting. Figure 9 identifies the various points in the network that can potentially limit system throughput, as analyzed in the W-Note, and it is of considerable interest to verify these issues as a step in proposing system improvements. For example, a recent experiment was carried out in which stream capacity was requested for four simultaneous 64-kbps full-duplex conversations over the channel at a burst rate of 772 kbps, and the PSAT refused the request. Sufficient stream capacity for three, but not four, conversations was successfully obtained. W-Note-33 had predicted that five such converations could be supported under these circumstances. Detailed discussion with the BBN PSAT team on the allocation of space in the PODA frame has allowed us to account for the difficulty. First, the current fixed-PODA (FPODA) implementation allows for growth to 10 stations (rather than accommodating only the current 4) and allots fixed reservation space in the FPODA frame for each of the 10. This consumes about 3,000 of the 16,384 bits in the PODA frame (at 772 kbps). 2,000 bits in each frame are reserved for datagram (not stream transmission). When all PODA and burst overhead is accounted for, the allowed stream capacity (including ST headers) is about 10,500 bits per frame, or 495 kbps. Four full-duplex PCM calls require 512 kbps, exclusive of ST headers. Hence, only three can be accommodated at the present time. At a future time to be determined, BBN has agreed to cut back on the reserved FPODA reservation and datagram space so that four calls can be accommodated. This will be a useful experiment in testing throughput limits in other parts of the system.

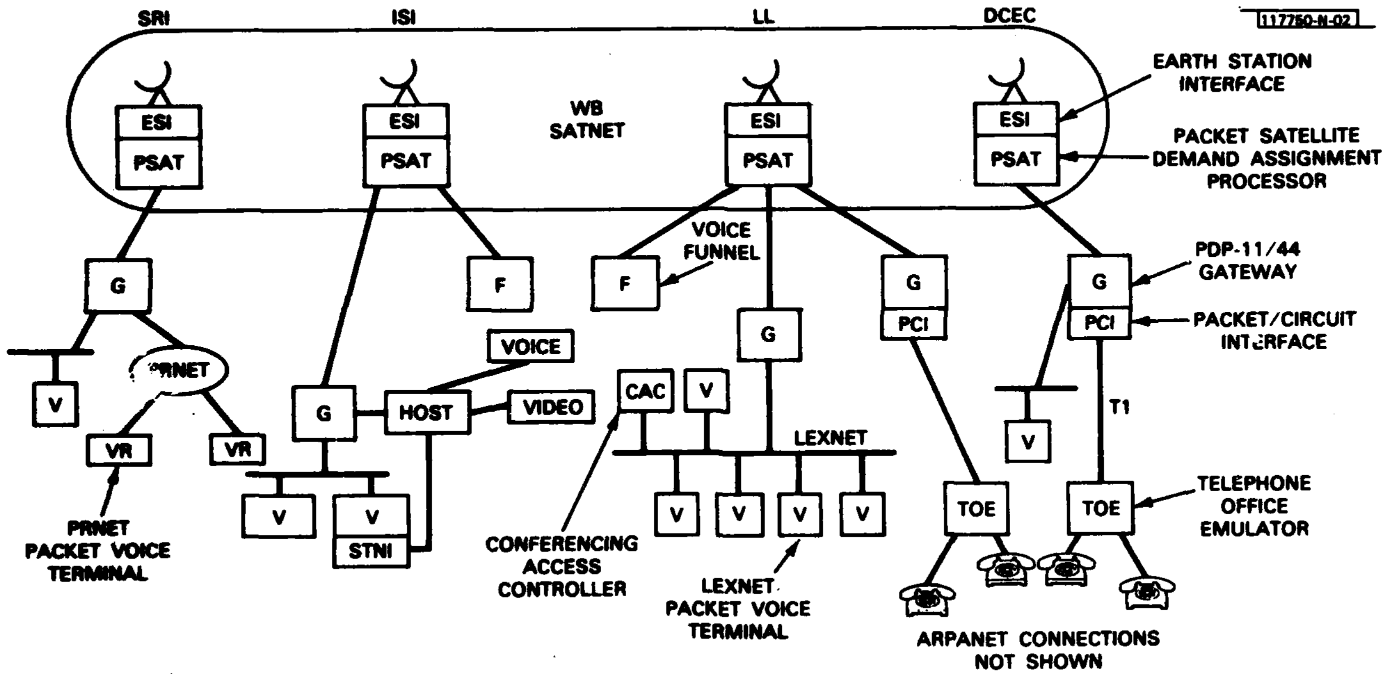

Fig. 7. Packet speech experiment system configuration as of June 1982.

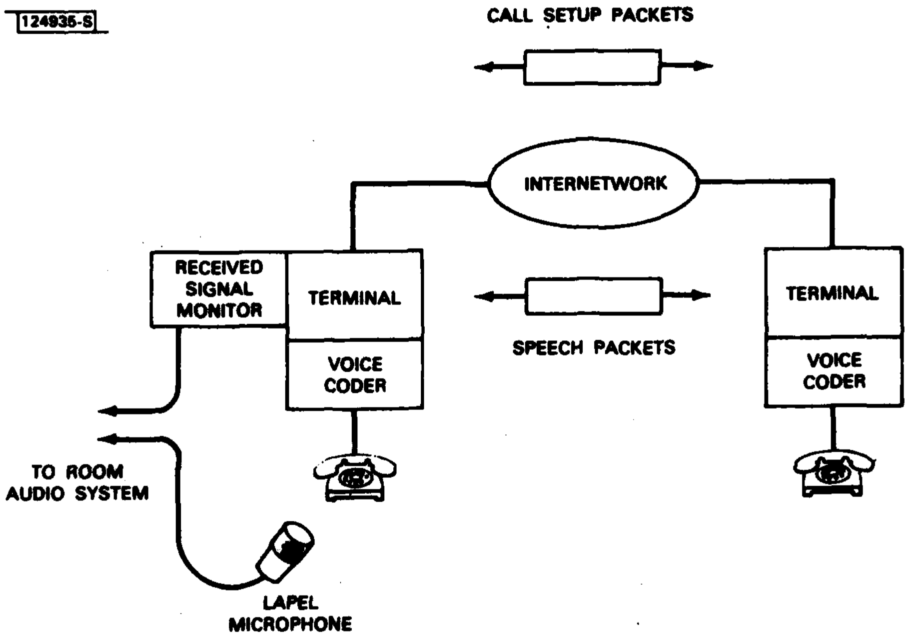

Fig. 8. Generic packet voice call demonstration setup.

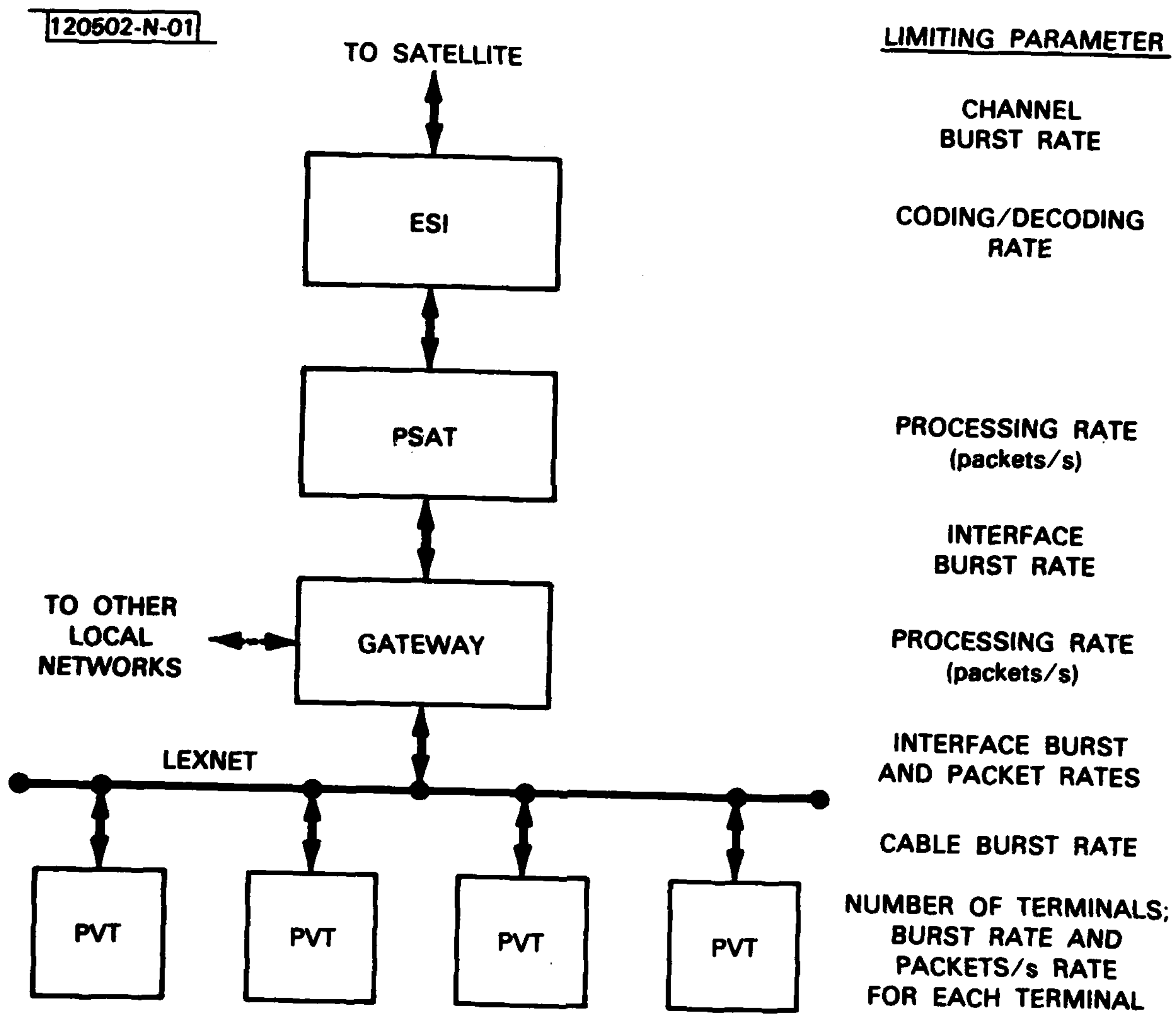

Fig. 9. *Wideband network site components and parameters limiting system throughput.*

VI. VOICE CONTROL OF NETWORK VOICE CONFERENCING

A. OVERVIEW

LPC and PCM conferences involving two to four participants have been successfully set up by voice between two LEXNETs connected through a gateway. A conference initiator first dials the Voice-Controlled Operator (VCOP) and names conference participants and the vocoder type during a voice dialogue with the VCOP. The VCOP then establishes the conference by ringing PVT telephones of conference participants. Upon answering, each participant is immediately part of the conference and can hear and talk to the current set of participants. A conference initiator can also call up the VCOP and train it to recognize his/her speech.

Successful setup of conferences followed extensive tests of the VCOP recognition accuracy and development of a human-VCOP dialogue controller. The recognition accuracy of the T580 word recognition system used in the VCOP was evaluated using PVTs on a local LEXNET and PCM, CVSD, and LPC speech coding. Recognition accuracy ranged from 85 to 98 percent correct. This accuracy can support human-VCOP voice interaction when talkers confirm recognition responses. A human-VCOP interaction protocol which performs well at these recognition rates was developed and tested with six male talkers. All successfully set up 10 conferences, five with 64-kbps PCM coding, and five with 2.4-kbps LPC coding. Conference setup times typically ranged from 40 to 70 s and times to train the VCOP to recognize each talker's words ranged from 4 to 6 min.

B. RECOGNITION ACCURACY OVER A LEXNET

The recognition accuracy of the T580 word recognition system in the VCOP was evaluated to determine whether it can support a conference setup dialogue with all types of speech coding and also to uncover any LEXNET or PVT system characteristics which need modification to improve recognition accuracy. Performance was evaluated using a 20-word vocabulary (the digits zero to nine, stop, start, yes, no go, help, erase, rubout, repeat, enter) four male talkers, and two female talkers. Ten repetitions of each word were used for training. Each talker read through the list of 20 words ten times (200 utterances) under each test condition. Tests were performed on a dedicated LEXNET with two PVTs. One was a normal PVT used by the talker and the other was a modified PVT which was part of the VCOP. Tests were performed over the LEXNET using (1) 64-kbps PCM coding without silence detection, (2) 16-kbps CVSD coding without silence detection, and (3) 2.4-kbps LPC coding with silence detection. In addition, recognition accuracy was measured for two talkers using a direct microphone input to the T580 recognition system and also using the three types of coding both with and without silence detection. A side tone was present in the talker's handset during all tests to help reduce variations in the input speech level to the speech coding cards and the recognition system.

Recognition accuracy was highest (99 percent correct) with direct microphone input. Accuracy dropped slightly with PCM coding (average = 94.5 percent, range = 88 to 99 percent), with CVSD coding (average = 92.5 percent, range = 86 to 98 percent) and with LPC coding (average = 91 percent, range = 86 to 95 percent). The addition of silence detection in the PVT reduced recognition accuracy by roughly 3.5 percentage points from an average of 97.5 percent to an average of 94 percent for all coding schemes for the two talkers on whom extensive measurements were made. Degraded performance with speech coding was expected because of the reduced analog bandwidth of the PVT telephone compared to the bandwidth of the close-talking microphone supplied with the recognition system (3500 vs 8000 Hz) and because of the information lost in the coded speech. The addition of silence detection degraded performance presumably because it prevented the complex word endpoint detection algorithm in the recognition system from functioning correctly. When silence detection was active this algorithm was effectively replaced by the simple silence detection algorithms used in the PVT.

C. PERFORMANCE OF HUMAN/VCOP INTERACTION DURING CONFERENCE SETUP

An effective VCOP human interaction must improve on the recognition accuracy of the T580 recognition system (86 to 99 percent) by (1) restricting the set of allowable recognition responses at each step of the dialogue and (2) forcing the caller to verify all recognition responses. Verification involves aural echoing of the T580 recognition responses using the VOTRAX synthesizer followed by a "YES" or "NO" judgment by the caller. A conference setup dialogue with no recognition errors that uses these techniques is illustrated in Fig. 10. Note that any response except "YES" and "NO" is echoed for verification. Although not shown, an incorrect recognition result with a "NO" verification causes the previous VCOP prompt to be repeated. In addition, a "HELP": response provides additional information about allowable responses and a "UH?", "WHAT?" or any other unrecognizable response causes the previous prompt to be repeated.

The above conference setup procedure was evaluated using six male talkers on a dedicated LEXNET with two PVTs (one for the talker, one part of the VCOP). The conference setup procedure was stopped as soon as the initiator-VCOP dialogue was complete. Each talker set up ten two-to-four member conferences. Half of these were set up using 64-kbps PCM coding without silence detection and half were set up using 2.4-kbps LPC coding with silence detection.

The vocabulary used during tests included twenty words: four commands (YES, NO, HELP, DONE), three vocoder types (PCM, LPC, CVSD), ten participant names (Lippman, Hsiung, Weinstein, Feldman, Gold, Blankenship, Forgie, O'Leary, Casner, Craighill) and three group names which describe grouping of three or more participants (24, network, speech). Talkers initially trained the recognition system by producing each word five times.

Training time for the 20-word vocabulary was typically four to six minutes. All talkers learned how to set up a conference with minimal instruction and conference setup items typically ranged from 40 to 70 s depending on the speech coding used and the number of participants. All conferences were set up successfully and the VCOP made no fatal mistakes.

The recognition accuracy averaged over all utterances produced by talkers during the setup procedure including YES/NO confirmations was 97.5 percent for PCM coding and 96.5 percent for LPC coding. These relatively high rates were caused by the restricted vocabulary size used at each stage of the human interaction.

D. VCOP/NVP INTERACTION

PVT software was modified to add new NVP tokens needed to set up conferences with the VCOP and to support communications between the PVT and PDP-11/44 which are parts of the VCOP. The software in the PVT that is part of the VCOP was modified to include a low-level protocol used to communicate with an 11/44 which is also part of the VCOP. This 11/44 controls the T580 recognition system and the VOTRAX synthesizer using RS-232 connections and communicates to a PVT over a RS-232 line. It answers the PVT phone when someone dials the VCOP, obtains conference setup information using the recognition system and the synthesizer, and then passes this information to the VCOP's PVT. The VCOP's PVT sends new Please-Join-A-Conference tokens to PVTs of all conference participants after receiving conference setup information. Participant PVTs ring and when answered, establish point-to-point connections with other conference participants after exchanging NVP tokens with the conference access controller. Participants then automatically join the conference as soon as they pick up their phones.

E. NETWORK TESTS

Voice conferences have been set up between participants on two LEXNETs at Lincoln connected via gateways and a PSAT. The VCOP and conference access controller were connected to one LEXNET along with one or two PVTs. One to three PVTs were on the other LEXNET. PCM and LPC conferences were established by voice by an initiator on either the VCOP's LEXNET or on the remote LEXNET. Conferences included from two to four participants. The human-VCOP interaction was always successful. At the end of this interaction

PVT telephones of all participants rang and upon answering, each participant immediately joined the conference. Some problems with the NVP protocol used to establish and terminate the conference were discovered. The NVP interaction is now being made more robust.

124841-N

INITIATOR DIALS VCOP NUMBER FOLLOWED BY USER ID NUMBER

VCOP "HELLO IS THIS MR. WEINSTEIN?"
CALLER "YES"
VCOP "WOULD YOU LIKE TO SET UP A CONFERENCE?"
CALLER "YES"
VCOP "VOCODER TYPE PLEASE"
CALLER "LPC"
VCOP "IS THAT LPC?"
CALLER "YES"
VCOP "PLEASE SAY PARTICIPANT OR GROUP NAME OR DONE"
CALLER "SPEECH"
VCOP "IS THAT SPEECH?"
CALLER "YES"
VCOP "NEXT"
CALLER "DONE"
VCOP "IS THAT DONE?"
CALLER "YES"
VCOP "THANK YOU, PLEASE WAIT"

PVT TELEPHONE OF EACH MEMBER IN THE GROUP NAMED "SPEECH" IS RUNG
MEMBER AUTOMATICALLY ENTERS CONFERENCE WHEN PHONE IS PICKED UP

Fig. 10. Sample dialogue between a conference call initiator and the VCOP. Note that the reply "speech" to the query "please say participant or group name or done" refers to a predefined group of participants interested in speech processing. Individual names could also have been entered.

REFERENCES

1. E. McLarnon, "A Method for Reducing the Frame Rate of a Channel Vocoder by using Frame Interpolation," Proc. 1978 Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), Washington, DC (April 1978), pp. 458-461

2. P.E. Blankenship and M.L. Malpass, "Frame-Fill Techniques for Reducing Vocoder Data Rates," Technical Report 556, Lincoln Laboratory, M.I.T. (26 February 1981), DTIC AD-A099395/6.

3. T.E. Tremain et al., "Implementation of Two Real Time Narrowband Speech Processing Algorithms," EASCON '78 Record, pp. 698-708 (September 1978).

4. P.E. Blankenship et al., "The Lincoln Digital Voice Terminal System," Technical Note 1975-53, Lincoln Laboratory, M.I.T. (25 August 1975), DDC AD-A017569/5.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: ESD-TR-82-155

2. GOVT ACCESSION NO.: AD-A126880

3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): Packet Speech Systems Technology

5. TYPE OF REPORT & PERIOD COVERED: Semiannual Technical Summary 1 April — 30 September 1982

6. PERFORMING ORG. REPORT NUMBER:

7. AUTHOR(s): Clifford J. Weinstein and Peter E. Blankenship

8. CONTRACT OR GRANT NUMBER(s): F19628-80-C-0002

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
Lincoln Laboratory, M.I.T.
P.O. Box 73
Lexington, MA 02173-0073

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:
ARPA Order 3673
Program Element Nos. 61101E/62708E
Project Nos. 2D10/2T10

11. CONTROLLING OFFICE NAME AND ADDRESS:
Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, VA 22209

12. REPORT DATE: 30 September 1982

13. NUMBER OF PAGES: 52

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
Electronic Systems Division
Hanscom AFB, MA 01731

15. SECURITY CLASS. (of this report): Unclassified

15a. DECLASSIFICATION DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

packet speech	ARPANET	LEXNET
network speech	SATNET	digital channel vocoder
voice conferencing		

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report describes work performed on the Packet Speech Systems Technology Program sponsored by the Information Processing Techniques Office of the Defense Advanced Research Projects Agency during the period 1 April through 30 September 1982.